Andrew L. Kun*

University of New Hampshire
Electrical and Computer Engineering
Durham, NH, USA

Hidde van der Meulen

University of New Hampshire
Electrical and Computer Engineering
Durham, NH, USA
and
University College Dublin
School of Computer Science
Dublin, Ireland

Christian P. Janssen

Utrecht University
Experimental Psychology &
Helmholtz Institute
Utrecht, The Netherlands

Calling while Driving Using Augmented Reality: Blessing or Curse?

Abstract

We report on an experiment on the distracting effects of in-car conversations through augmented-reality glasses. Previous research showed that in-car phone conversations can be distracting, but that the distraction might be reduced if the remote caller receives visual information about the driving context. However, what happens if such video sharing becomes bidirectional? The recent introduction of commercial augmented-reality glasses in particular might allow drivers to engage in video-supported conversations while driving. We investigate how distracting such video-based conversations are in an experiment. Our participants operated a simulated vehicle, while also playing a conversational game (Taboo) with a remote conversant. The driver either only heard the remote conversant (speech-only condition), or was also able to see the remote person in a virtual window that was presented through augmented reality (video call condition). Results show that our participants did not spend time looking at the video of the remote conversant. We hypothesize that this was due to the fact that in our experiment participants had to turn their head to get a full view of the virtual window. Our results imply that we need further studies on the effects of augmented reality on the visual attention of the driver, before the technology is used on the road.

1 Introduction

Augmented reality combines real and virtual objects in a real environment (Azuma, Baillot, Behringer, Feiner, Julier, & MacIntyre, 2001). It has the potential to support driving-related activities, as reported throughout this special issue. However, as with any technology, it can be reappropriated in unanticipated ways. In a time-sensitive, safety-critical context such as driving, this has the potential to lead to dangerous behavior.

In this article, we report on an experiment to assess whether augmented reality-based video conversations between car drivers and remote callers cause additional visual distraction, compared to a situation in which only audio is shared. As it turns out, augmented reality has some potential in this area, which might make it a blessing. However, there are also pitfalls that might make it a curse. We also compare our findings to results from a previously reported experiment using conventional (nonaugmented) video-based conversations (Kun & Medenica, 2012). Taken together, the results from the two experiments help

*Correspondence to andrew.kun@unh.edu.

us to identify the blessings and the curses of augmented reality-based video conversations in cars.

Before discussing our study in more detail, we will review related work on the use of augmented reality while driving, using it during automated driving, and on driving and talking.

2 Related Work

2.1 Augmented Reality and Driving

Milgram and Kishino (1994) define augmented reality as the case when the real world is augmented with displays of computer-generated (virtual) objects. In their taxonomy, the concept of augmented reality is a subset of the broader concept of mixed reality, where mixed reality refers to any case when the real world and computer-generated content are merged for the user. The relative amount of real-world and virtual content is different for different cases of mixed reality. However, mixed reality refers to all of the cases that fall between two extremes: showing the unmodified real environment (no virtual content), and showing a virtual environment (no tie to the real world, commonly designated as virtual reality or VR).

Traditionally, augmented-reality devices have been constrained to experiments in laboratories. However, recent technological advances have resulted in broader availability of augmented-reality devices, including the Microsoft HoloLens, which is the device we used in the experiment presented in this article.

Given the fact that augmented-reality devices are becoming more broadly available, we expect that in the not-so-distant future they might be used in many domains. One domain where augmented-reality devices might be encountered is vehicles and transportation. Gabbard, Fitch, and Kim (2014) provide a broad overview of opportunities and challenges of introducing augmented-reality displays in vehicles. They argue that augmented-reality devices likely need two characteristics to be useful in providing task support, such as navigation instructions. First, they should be see-through devices, because this would allow drivers to see all of the relevant

real-world information, such as changes in road geometry and the position of other road users. Second, they should be able to accurately position virtual objects with respect to the real world, which would allow them to provide meaningful support with respect to actions that are related to the real world. Gabbard and colleagues also argue that much work is needed to understand how such devices will influence drivers. We agree, and add that this work is not only needed in the realm of tasks which directly support the driving task (such as for navigation and traffic warnings), but also for cases when the augmented-reality device is appropriated to perform non-driving-related tasks.

Smith, Gabbard, Burnett, and Doutcheva (2017) present results that can get us closer to understanding the effects of augmented-reality devices on driving. In a driving simulator study they compared participants' visual behavior for two displays: a see-through, head-up augmented-reality display that required small changes in gaze angle away from the road, and a head-down LCD that required larger changes in gaze angle. Smith and colleagues found that participants cast longer glances at the augmented-reality device than at the head-down device, indicating perhaps that they were more comfortable taking their eyes off the road for long periods of time in the case when the required change in gaze angle was smaller.

Much of the work on exploring augmented reality and driving was conducted in the domain of navigation. In our own prior work, we found that augmented reality could be beneficial in navigation, allowing drivers to keep their eyes on the road instead of needing to cast glances at an in-vehicle map (Medenica, Kun, Paek, & Palinko, 2011). We simulated augmented reality by presenting augmented reality-like instructions on the simulator's screens. Bolton, Burnett, and Large (2015) explored the effects of different ways to present navigation information using augmented reality on the ability of drivers to follow the instructions. They found that augmented-reality boxes that enclose a landmark improved performance on the navigation task, compared to more traditional approaches to providing instructions, such as displaying arrows to indicate turns. Fröhlich, Baldauf, Hagen, Suette, Schabus, and Kun (2011)

explored an augmented reality-like display that overlaid navigation instructions on the live feed of the road on a head-down display. However, they found that, while drivers valued visual information, they preferred a conventional map display over the head-down augmented-reality display. Bark, Tran, Fujimura, and Ng-Thow-Hing (2014) experimented with a laboratory-based volumetric head-up display, which is capable of presenting 3D information. They argue that for navigation information, where depth perception can be important, augmented-reality information should be presented in 3D.

In contrast to these efforts, the work in this article focuses on communication with a remote conversant. This is a relevant issue, since communicating with remote conversants is common while driving, and it is an activity that can even lead to crashes (e.g., see the findings of Neale, Dingus, Klauer, Sudweeks, and Goodman [2005] in the 100-car study). Furthermore, communication while driving is an integral part of using vehicles for some professionals, such as first responders (Kun, Wachtel, Miller, Son, & Lavallière, 2015).

2.2 Augmented Reality in Automated Cars

Augmented-reality devices can also play a role in (semi-)automated, or (semi-)autonomous vehicles, especially for the levels of automation where the human still has a supervisory role (e.g., up to Society of Automotive Engineers level 3; SAE, 2015). For example, augmented-reality devices can augment the driving experience by highlighting other traffic and objects (e.g., objects that the car is unsure about), and can thereby provide a tool for designers to increase user trust in automated vehicles (Wintersberger, Sawitzky, Frison, & Riener, 2017). On the other hand, augmented-reality devices can allow drivers to transform vehicles into places for productivity and play (Kun, Boll, & Schmidt, 2016).

However, using augmented-reality devices in automated vehicles, where users might look away from the road as they use the device, could result in motion sickness. More research is needed to understand how we can reduce this negative effect within augmented reality (note that motion sickness can be very strong when virtual reality is used in vehicles) (McGill, Ng, & Brewster, 2017).

The current article addresses the use of augmented reality in manually driven (i.e., nonautomated) vehicles. As we argued in the preceding text, using augmented-reality technology in manually driven vehicles might prove to be a blessing or a curse. Of the curses, this might include distraction from driving. However, if we observe distraction, this might not be limited to nonautomated settings. Meta-reviews suggest that drivers distract themselves with non-driving tasks as automation in the car increases (De Winter et al., 2014). In particular, at automation levels where control of the vehicle is shared between the driver and the vehicle (i.e., SAE levels 1 to 3; SAE, 2015), secondary tasks might keep visual attention away from the road, and engagement with other tasks might harm driving performance. Therefore, even though our study is tested in nonautomated driving, we expect that our results will be relevant to automated driving systems.

2.3 Driving and Talking

The particular task that we consider in our experiment is that of a conversation with a remote person. Such practices have been observed in naturalistic driving studies (e.g., Dingus et al., 2016; Klauer et al., 2014). In general, it has long been known that performing additional tasks while driving can distract attentional resources away from driving (e.g., see the work of Senders, Kristofferson, Levison, Dietrich, and Ward [1967]). Phone calls in particular are a source of distraction for three reasons. A first source of distraction is the visual-manual operation of the phone, which prevents the driver from looking continuously at the road (Brookhuis, de Vries, & de Waard, 1991; Salvucci, 2001). However, modern interfaces have been designed to reduce this type of interference (e.g., hands-free interfaces).

A second source of distraction is the conversation itself: talking and thinking about what to say can distract, even when your eyes stay on the road (Iqbal, Ju,

& Horvitz, 2010; Kunar, Carter, Cohen, & Horowitz, 2008; Strayer & Johnston, 2001). Pfleging, Schneegass, and Schmidt (2013) therefore explored the idea of sharing video from the road, which might discourage placing a call to a driver. In non-driving domains, technology has also been designed to inform callers about situations where it might not be appropriate to call (Grandhi & Jones, 2015), or to help them postpone a call (Böhmer, Lander, Gehring, Brumby, & Krüger, 2014).

A third source of distraction is the lack of shared context between the driver and the remote caller. Compared to holding a conversation with a passenger, a remote caller might not notice when the driver is overloaded by traffic demands and might continue to be engaged in a deep conversation. Research has therefore looked into how sharing the driver's context in various ways, including via additional audio and video, can reduce distraction (Charlton, 2009; Gaspar et al., 2014; Janssen, Iqbal, & Ju, 2014; Maciej, Nitsch, & Vollrath, 2011; Schneider & Kiesler, 2005). In general, sharing aspects of the driver's context with a remote caller through video seems to reduce the negative impact that the conversation has on driving safety (Gaspar et al., 2014).

However, such context sharing might work both ways. In particular, when a driver shares video with a remote caller, technology might also share video from the remote caller with the driver. Since video-sharing results in the driver receiving more information (i.e., not just audio, but also visual stimulation), there is a risk of increased distraction and mental load. To assess how strong such distraction is, Kun and Medenica (2012) explored the effects of this two-way video calling while driving in a driving simulator study. The study contrasted driving under two conditions: speech-only and video call. In the speech-only condition the driver and the remote conversant could hear each other, but not see each other. In the video call condition, they could also see each other. For the driver this was through a video screen that was fixed to the top of the center console of the car. We found that in the video call condition drivers looked away from the road ahead more often than in the speech-only condition. This effect was confined to driving on straight roads. On curvy roads, drivers' attention to the road ahead was not affected by

the type of call (speech-only or video call). However, even straight roads require visual attention. We concluded at the time that video calling presents a problem, because it is likely that drivers will—sometimes incorrectly—assume that some road segments are safe enough to engage in video calling, which in turn will reduce their visual attention to the road ahead.

2.4 Conversation Tasks in Driving Studies

The context in which we used augmented reality was a conversation task. A variety of conversation and language-based tasks have been used in driving studies, ranging from repeating words (Van der Heiden et al., 2018), to question answering (e.g., Iqbal et al., 2011), to conversation games (e.g., Crundall et al., 2005; Kun, Shyrokov, & Heeman, 2010), to relatively structured conversation for example about predefined topics (e.g., Drews, Pasupathi, & Strayer 2008; Gaspar et al., 2014; Janssen et al., 2014; Schneider & Kiesler, 2005), to free conversation (e.g., Charlton, 2009; Maciej et al., 2011).

Our aim was to simulate a natural conversation, while also exerting sufficient experimental control to avoid confounds in measurement. To accomplish this, it was important to select a spoken task that is (1) engaging and (2) in which both the driver and the other conversant have to speak (Kun, Shyrokov, & Heeman, 2010). Some spoken tasks that meet these requirements are the parallel 20 questions game (Kun, Shyrokov, & Heeman, 2013), the game of Taboo (Heeman, Meshorer, Kun, Palinko, & Medenica, 2013; Kun et al., 2012; Kun, Palinko, Medenica, & Heeman, 2013), and collaboratively generating fictional stories (Janssen et al., 2014). Note that even further controlled tasks such as question answering (e.g., Iqbal et al., 2011), might not be engaging enough, whereas free conversations (e.g., Charlton, 2009; Maciej et al., 2011) might result in unbalanced conversations in which drivers vary in how much they only listen or talk.

Taboo is the task we used in the current study, and it is described in the Method section. An additional benefit of Taboo is that it makes it easier to compare our new

Figure 1. *Participants operated a simulated vehicle and wore a HoloLens augmented reality device. They used HoloLens to communicate with a remote conversant.*

results to those from our prior work on in-car two-way video calling (Kun et al., 2012), because the prior study also used Taboo.

3   Method

3.1 Participants

Fourteen participants were invited to the experiment. We discarded data from four participants; three participants lacked sufficient English language skills to complete the spoken task, and for one participant we encountered technical problems during data collection. The remaining 10 participants (2 female; 8 male) ranged between 19 and 23 years of age ($M = 19.6$ years, $SD = 1.3$ years). The reduced number of participants harms the power of our experiment, and results should therefore be interpreted with caution.

All participants were students and they received course credit for their participation. The experiment was approved by the UNH Institutional Review Board for the Protection of Human Subjects in Research. Informed consent was obtained from all participants.

3.2 Tasks and Equipment

Participants engaged in two tasks: driving and a spoken task (game of Taboo).

3.2.1 Driving Task. In the driving task, participants drove at a fixed speed of 50 mph behind a yellow lead car. The driving environment was a straight rural road with two lanes. There was no other traffic apart from the lead vehicle.

The hardware consisted of the DriveSafety simulator (see Figure 1). This can be considered a high-fidelity simulator for three reasons. First, the participant sits in an actual car frame. Second, the simulator has a moving base, which can tilt forward and back; this allows participants to experience the motion of acceleration, deceleration, and bumps in the road. Third, participants have a 180° field of view, generated by three projectors, and

they can also view simulated rear- and side-view mirrors implemented through LCD displays.

3.2.2 Spoken Task.

For the spoken task, participants played a game of Taboo, which we also used in previous driving-related experiments (Heeman et al., 2013; Kun et al., 2012; Kun, Palinko, et al., 2013). The aim of the game of Taboo is to guess a target word, which is usually presented on a card. To this end, one player (the describer) describes the target word (e.g., soccer), but is not allowed to use the target word, and five predefined Taboo words in their description (e.g., ball, game, grass). These Taboo words are also presented on the same card. The other player (the guesser) keeps guessing the word until they find it, or until some time limit has passed, or until they give up on guessing the word.

In our implementation the driver was always the guesser, and the remote conversant was the describer. We implemented a 1-minute time limit on guessing a given word. The Taboo cards were discussed in the same order for each participant.

The role of the remote conversant was always played by the (same) experiment leader. We chose this configuration to limit the variation of the driver's cognitive load. Cognitive load variation was reduced by the fact that the experiment leader practiced the word descriptions ahead of time, which minimized the occurance of disfluencies, and outright mistakes, in his speech. Also, the driver did not need to change between different game roles (describer and guesser), which otherwise could have contributed to cognitive load variations during the experiment (cf. Kun, Shyrokov, et al., 2013).

For the hardware, we used the Microsoft Hololens[1] (shown in the driving simulator in Figure 1). The device was worn on top of a Pupil Labs[2] headworn eye-tracker (see Figure 2). The eye tracker was controlled by software that ran on a Dell XPS 13 Windows 10 laptop. We also ran Skype on the same laptop for the remote conversant (played by the experiment leader). In our (augmented) video call condition, participants could see each other via a Skype window; the window was presented
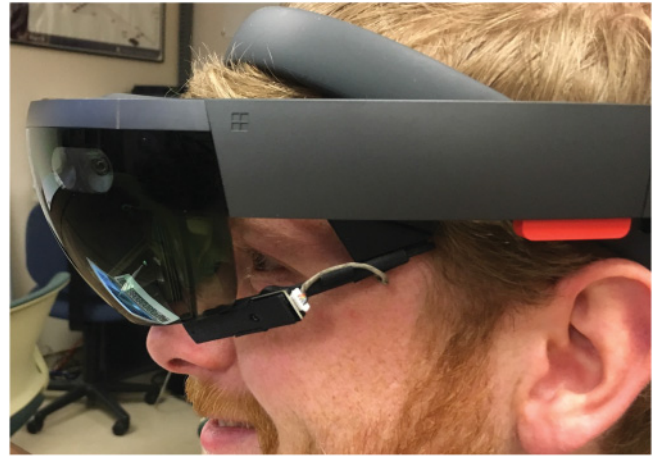

Figure 2. *Participants wore HoloLens while driving. The Pupil Labs eye tracker fits under HoloLens. In this picture we can see that there is ample space under HoloLens for the tracker's left eye camera.*

through the Microsoft Hololens to the driver (see Figure 3), and through the laptop to the experiment leader.

The HoloLens can project visual stimuli in a field of view of 40° wide by 20° high, relative to the viewer. It allows users (or programmers) to pin virtual objects to specific locations in the real world. In our case, we instructed drivers to pin the Skype conversation window to the top of the center console (Figure 3), as this was similar to the location where a video display was used in our previous study with conventional video (Kun et al., 2012). Thus, to see the Skype window, drivers had to look to the right and slightly down. If their eyes were on the road, they would not see the window in their peripheral vision (or at least not the entire Skype window) due to the relatively narrow field of view of the HoloLens. This is in contrast to the prior experiment in which the physical video screen was visible in peripheral vision. The HoloLens generates directional sound, which gave the driver the impression that the audio also came from the pinned window.

In the speech-only version, participants still wore the HoloLens, but could hear the remote caller only via the speakers of the HoloLens. At the location of the Skype window they saw only a Skype logo, if they looked there.

---

[1]https://www.microsoft.com/en-us/hololens

[2]https://pupil-labs.com

Figure 3. *The Skype window displayed by HoloLens is shown in the center of this image. Skype was pinned to a location on top of the center console. The world camera of the eye tracker fits under HoloLens, and was used to take this picture.*

3.3 Design

A one-factor within-subjects design was used. The factor was conversation mode, with two levels: speech-only or (augmented) video call. The order of the conditions was counterbalanced between participants.

3.4 Procedure

Upon arrival, participants signed a consent form and received a general explanation of the procedure, including how to wear the eye-tracker and HoloLens. They then read an introduction to Taboo and practiced the Taboo game with the experiment leader until they felt confident playing. This was followed by a practice drive in the simulator, again until the participant felt comfortable with the task (typically only a couple of minutes).

After the practice, the experiment leader placed the eye-tracker and the HoloLens upon the participant's head. The devices were adjusted to fit comfortably, and such that the eye-tracker could both track the participant's eyes and such that its world camera could provide a view through the HoloLens. Once the adjustments were made, participants were asked not to touch the devices. Then the Pupil Labs manual marker calibration procedure was completed to calibrate the eye-tracker. Next, the HoloLens and Skype application were started. Participants had to pin the Skype window to the top of the center console, just underneath the windscreen (Figure 3).

After these steps the experiment leader moved to another room and the remaining communication was via Skype. Participants completed the two experimental conditions, driving while conversing in a speech-only and an (augmented) video call condition. For each

condition, there was one practice session in which participants operated the vehicle while also guessing 8 Taboo cards. After the practice, participants started a new driving scenario, and the eye-tracking and driving data was logged. Roughly twenty seconds into the drive, the first Taboo card was started, and in total 20 Taboo cards were played during each data collection scenario.

At the end of the experiment, a questionnaire was completed via Lime Survey. The questionnaire collected demographic data as well as the participants' views on driving with HoloLens and their experience of the two experimental conditions, speech-only and (augmented) video call. In total, the experiment lasted approximately 50 minutes.

3.5 Measures

We tracked participant gazes during the experiment using a Pupil Labs eye-tracker that fits underneath the HoloLens. All measures were obtained for each participant and interaction type (speech-only and video call), and then averaged over all participants. We collected the following measures:

- Percent dwell time (PDT) on the road ahead (i.e., percentage of time drivers spent looking at the forward road). This measure was calculated using eye-gaze data obtained from the Pupil Labs eye-tracker. The Pupil Labs software uses two eye cameras to determine the user's gaze location relative to the eye-tracker itself. It also uses a view of the environment, obtained through a so-called world-camera, to determine gaze location on areas of interest in the environment. This latter calculation is performed by observing the location of 2D barcodes in the environment. We placed 2D barcodes in the driving simulator, and the software used them to mark all gaze instances as either being directed at the (simulated) road ahead, or at other areas in the driver's field of view. Decreased PDT on the road indicates reduced visual attention.
- Standard deviation of lane position (SDLP), as defined in SAE J2944 (SAE, 2015). We collected lateral lane position data from the driving simu-

lator. Increased SDLP can indicate worse driving performance.
- Number of missed cards in Taboo. A large number of missed cards might indicate either that the combination of the driving and spoken tasks was too difficult for the participants to complete, or that participants did not pay attention to the Taboo game.
- Levels of agreement with preferential statements on 5-point Likert-type questions.

We calculated PDT and SDLP over 3 minute-long segments that started 20 seconds after the beginning of an experiment. We did this regardless of how long it took to complete the 20 Taboo cards for an experiment. Eye-tracker data was collected at 30 Hz, while driving-simulator data was collected at 10 Hz. We counted the number of missed cards over the total of 20 cards played by each participant over the two conversation modes. We asked participants to indicate their level of agreement with Likert-type questions at the end of the experiment, after they completed the Taboo game with both conversation modes.

Additionally, at the end of the experiment, we solicited written comments about the experiment from participants.

4 Results

4.1 Performance Data: Visual Attention, Driving, and Game Performance

For technical reasons we had to exclude eye-tracking data for three participants. For one participant the HoloLens covered the world camera; thus we have no way to establish where this participant directed a gaze during the experiments. For two other participants the data collected by the Pupil Labs eye-tracker appears to have been of poor quality, as the Pupil Labs software reported that the confidence of gaze tracking was below our 70% confidence threshold for most of the experiment. We compared the PDT values for the remaining 7 participants using a paired t-test. We
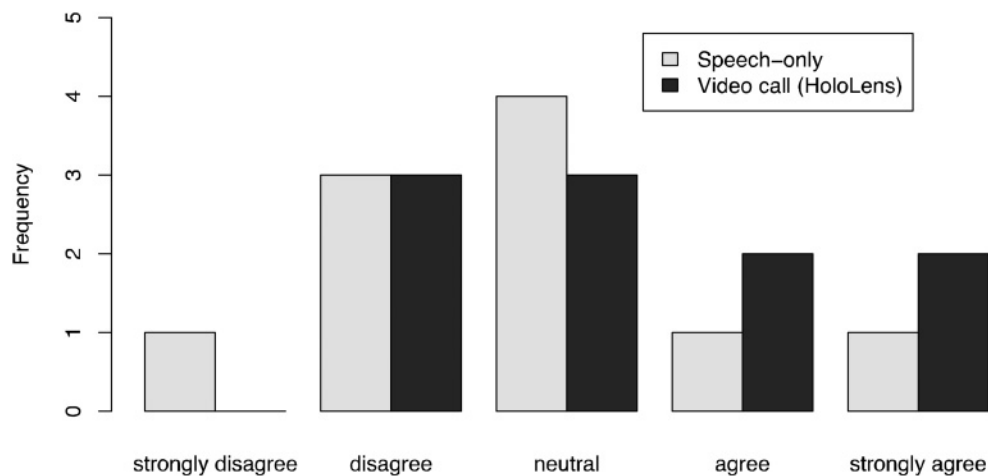
Figure 4. *Participants' subjective reports on how distracted they felt from driving when they were in the speech-only or video call condition.*

did not observe a significant difference between the speech-only ($M = 95.7\%$, $SD = 3.2\%$) and video call ($M = 96.8\%$, $SD = 2.8\%$) conditions, t(6) $= -1.525$, $p = .178$. The high magnitude of the PDT values is comparable to the previous study with a physical video screen (Kun et al., 2012). However, in contrast to the current study, the previous work found a difference between speech-only and video call (on straight roads).

A paired *t*-test for all 10 participants did not reveal any differences in SDLP for the speech-only ($M = 0.21$ m, $SD = 0.09$ m) and video call ($M = 0.22$ m, $SD = 0.12$ m) conditions, t(9) $= -.879$, $p = .402$. This is similar to the findings in Kun et al. (2012), and not surprising, since participants spent roughly the same amount of time looking at the road ahead in the two conditions.

Finally, participants successfully guessed most of the 20 Taboo cards. The number of words they could not guess was low for both the speech-only ($M = 1.5$, $SD = 1.4$), and the video call ($M = 1.8$, $SD = 1.8$) condition.

4.2 Subjective Experience Data: Preferential Statements and Written Feedback

We wanted to assess the viability of the speech-only and video call modes of interaction in real driv-

ing. For this reason, we asked participants to provide responses to statements that related speech-only and video call phone conversation to their own vehicles. Figure 4 shows the distribution of drivers' response in agreement to the question "In the experiment I was distracted from driving when using" speech-only (grey bars), or the video call (black bars). The distributions strongly overlap, suggesting that both forms were experienced as equally distracting.

We also asked participants to give a forced-choice reply to the question: "Which phone conversation distracted you more from driving?" Fully eight participants selected the video call and two selected speech-only. The written comments of two participants gave insight why the video call might not always have been distracting. Participant 1 wrote: "because the video was outside my field of view, it did not add any distraction to my driving abilities." Participant 6 wrote: "Hololens is cool, but the placement of the video display may have been better off closer to my field of view of the road so I didn't have to look so far if I wanted to see the video."

Finally, we also asked participants whether they would engage in speech-only calls and video calls in their own car. Figure 5 shows the distribution of results. Seven participants strongly agreed, or agreed, that they would engage in speech-only conversations (e.g., phone calls) while driving in their own car. However, only two
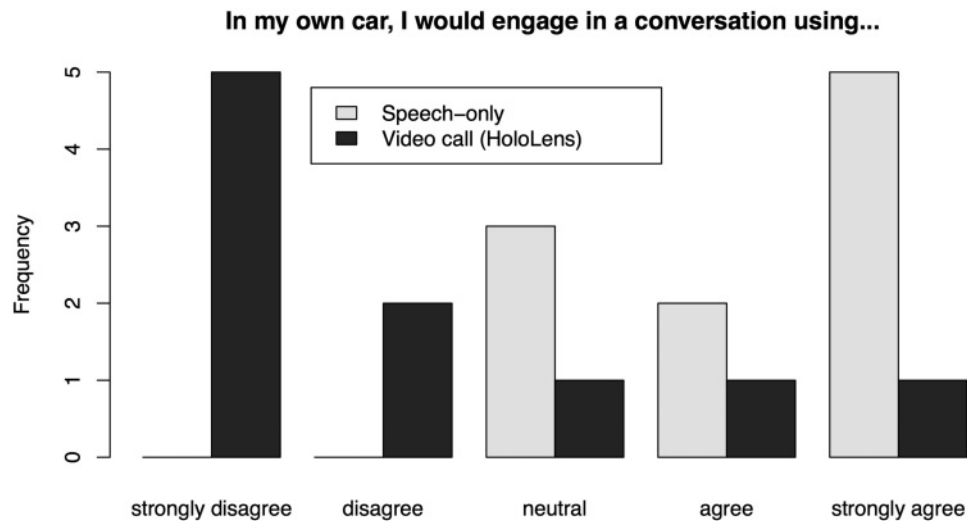
In my own car, I would engage in a conversation using...

Figure 5. *Subjective reports whether users would engage in a speech-only or HoloLens-based conversation when driving in their own car.*

participants said the same for video calls. We compared these levels of agreement for each participant, and found that eight participants rated the likelihood of engaging in a video call in their own car lower compared to engaging in a speech call (one participant gave them equal likelihood of "strongly agree"; one participant rated the video call slightly higher with "agree").

Importantly, the number of participants who would consider engaging in this practice using video call *is not 0*—two participants were enthusiastic about using video call in the car. This is also reflected in this comment from participant 9 (who "agreed" to the statement about using HoloLens in the car): "It was a nice experience wearing such an advanced tech." The two participants who rated it most likely to use video call in their own car also provide the lowest ratings to the question of how distracting they experienced the video call to be.

5 General Discussion

5.1 Augmented Reality: A Blessing or a Curse?

Our results show that some people are willing to use augmented reality in the car to aid conversation with remote callers (see Figure 5). We investigated whether augmented-reality devices are a blessing or a curse when used for such conversations. We found evidence for both. Our starting point in the research was the trend that video sharing from a driver to a remote caller has benefits (Gaspar et al., 2014), but that such sharing might also be inversed, where the driver might want to look at video information shared by the remote caller. When presented on a real in-car display, such screens can distract (Kun et al., 2012). In theory, augmented reality can be a blessing in that the video screen can be presented at other locations, including close to traffic views. However, it might also form a curse, in that the video images might take visual attention away from the traffic environment.

In our experiment, we compared a situation where a driver interacted with a remote conversant using audio only, with another situation where there was also an augmented-reality window that presented a video of the remote conversant. We found no difference in the percent dwell time of the road (a proxy for distraction) between the two conditions. This contrasts with prior work in which video was presented on an in-car physical screen (Kun et al., 2012), where drivers looked away from the road more when the screen was used.

Our interpretation of this difference between the two studies is that this is due to the presence or absence of peripheral stimuli and the associated action to see it. The in-car screen was visible in the periphery of the field of view of the drivers, if they were fixating locations on the road. By contrast, in our augmented-reality glasses, the field of view in which the HoloLens could project was relatively small (40° horizontally; 20° vertically). In order to see the Skype window then, participants had to make a head movement. In effect, this might have contributed to fewer glances to the screen in two ways. First, there were no distracting stimuli visible in periphery. Second, the need of making a head movement, and the associated energy costs, might prevent users from making such movements too often. This is consistent with findings on visual working memory, in which the requirements of head movements changed the way users interact with task interfaces (Ballard, Hayhoe, & Pelz, 1995) (see Gray, Sims, Fu, & Schoelles, 2006 for a more general interpretation of how the costs of information access influence decision making and interaction with an interface). The subjective data also confirm that participants did not like making a head movement; this is supported by the quotes from participants 1 (". . .video outside my field of view. . .") and 6 (". . .placement of the video display may have been better . . .[if]I didn't have to look so far. . ."), discussed above.

Although it is tempting to interpret this result as a blessing, it might still be a hidden curse. First, compared to just an eye movement, when a head movement is made, gaze position can be further removed from the road, and it also takes longer to return the gaze to the road, as a head movement is again required. Second, although we constrained the location of the Skype window to the center column, a future implementation of the technology might not prevent users from placing augmented-reality content at other locations, including those prominently in front of them. Therefore, users might block their own field of view with augmented-reality content. An interesting consideration in this regard is that eight of our drivers indicated in a post-experiment questionnaire that they did not think that the use of augmented reality distracted them from driving. It is an open question whether drivers also experience it in this way outside of the lab and when using the technology in a wider variety of driving environments.

5.2 Limitations

Within our experiment two types of limitations can be distinguished. The first set relates to the use of HoloLens, and how the use in the experiment might scale to a use outside of the experiment in actual cars, and as the technology and capabilities of augmented reality techniques develop. Most importantly, the HoloLens has a limited field of view. Combined with our requirement to place the Skype window on the center console (cf. Kun et al., 2012), this meant that the video screen was outside of the users' peripheral view and required a head movement to see. The need of a head movement can influence behavior in subtle ways (cf. Ballard et al., 1995). But, if the field of view of the augmented-reality glasses is expanded, our findings might not hold. For example, Xiao and Benko (2016) combined the HoloLens with peripherally placed LEDs, to augment the field of view. It is an empirical question whether use of such LEDs increases or decreases the distracting effect of HoloLens when used for conversations.

The use of the HoloLens device was new to our participants. Therefore, it is unclear how the use of this technique might develop with longer exposure. A particular relevant question is whether participants might start to feel more comfortable to make a head movement, once they know how the technology works.

For the Taboo game, the driver played the game with the experiment leader. As this was someone who was less familiar to them, they might have interacted differently compared to when speaking to someone whom they know well. One hypothesis is that participants might look more at the video screen when talking to a close friend or family member, resulting in lower PDT at the road ahead than in the current experiment. However, data is needed to confirm this.

Finally, there are two limitations related to the setup of the experiment itself. First, within the experiment

we asked participants to place the Skype window themselves. This might have led to slight variations in the position of the window, and thus it is possible that different users needed to turn their heads more or less to view the Skype window.

Second, we measured only a small set of ten participants, and of these only the eye-gaze data of seven could be used. This limits the power of the study. However, despite the low numbers, the observed results were very consistent, as evident in the small standard deviation of the PDT measure (2.8% and 3.2%, respectively, for the video and audio conditions). Moreover, we visually inspected the world camera video recordings for the two participants for whom the eye-gaze data was not available, and we looked for segments that show head motions towards the center console, as these would indicate possible glances at the Skype window. However, the videos indicate that the participants made only very few such head motions, which leads us to conclude that they made very few glances towards the Skype window.

6 Conclusion

Augmented-reality devices are perhaps a blessing, but they might turn out to be a curse. In our study, participants experienced little visual distraction from a video call that was played on the augmented-reality device. However, if drivers decide to look at such a video, they might either spend long periods of time looking away from the road as they turn their head towards a virtual window, or they might decide to pin the virtual window in a position where it overlaps the view of the road. Both of these scenarios might turn out to be distracting drivers from the primary task of driving. Thus, more work is needed to understand the conditions under which the blessing turns into a curse, particularly as at least some users seem willing to use augmented reality for in-car conversations. In other words, further work is needed to understand the effects of using augmented reality in cars on the drivers' visual behavior, and the resulting effects on their driving performance, such as in the work of Smith et al. (2017).

References

Azuma, R., Baillot, Y., Behringer, R., Feiner, S., Julier, S., & MacIntyre, B. (2001). Recent advances in augmented reality. *IEEE Computer Graphics and Applications, 21*(6), 34–47.

Ballard, D. H., Hayhoe, M. M., & Pelz, J. B. (1995). Memory representations in natural tasks. *Journal of Cognitive Neuroscience, 7*(1), 66–80.

Bark, K., Tran, C., Fujimura, K., & Ng-Thow-Hing, V. (2014). *Personal Navi: Benefits of an Augmented Reality Navigational Aid Using a See-Thru 3D Volumetric HUD.* Paper presented at the 6th International Conference on Automotive User Interfaces and Interactive Vehicular Applications.

Böhmer, M., Lander, C., Gehring, S., Brumby, D. P., & Krüger, A. (2014). *Interrupted by a Phone Call: Exploring Designs for Lowering the Impact of Call Notifications for Smartphone Users.* Paper presented at the SIGCHI Conference on Human Factors in Computing Systems.

Bolton, A., Burnett, G., & Large, D. R. (2015). *An Investigation of Augmented Reality Presentations of Landmark-Based Navigation Using a Head-Up Display.* Paper presented at the 7th International Conference on Automotive User Interfaces and Interactive Vehicular Applications.

Brookhuis, K. A., de Vries, G., & de Waard, D. (1991). The effects of mobile telephoning on driving performance. *Accident Analysis & Prevention, 23*(4), 309–316.

Charlton, S. G. (2009). Driving while conversing: Cell phones that distract and passengers who react. *Accident Analysis & Prevention, 41*(1), 160–173.

Crundall, D., Bains, M., Chapman, P., & Underwood, G. (2005). Regulating conversation during driving: A problem for mobile telephones? *Transportation Research Part F: Traffic Psychology and Behaviour, 8*(3), 197–211.

de Winter, J., Happee, R., Martens, M. H., & Stanton, N. A. (2014). Effects of adaptive cruise control and highly automated driving on workload and situation awareness: A review of the empirical evidence. *Transportation Research Part F: Traffic Psychology and Behaviour, 27*, 196–217. Retrieved from http://doi.org/10.1016/j.trf.2014.06.016

Dingus, T. A. Guo, F., Lee, S., Antin, J. F., Perez, M., Buchanan-King, M., & Hankey, J. (2016). Driver crash risk factors and prevalence evaluation using naturalistic driving data. *Proceedings of the National Academy of Sciences*, 201513271. Retrieved from http://doi.org/10.1073/pnas.1513271113

Drews, F. A., Pasupathi, M., & Strayer, D. L. (2008). Passenger and cell phone conversations in simulated driving. *Journal of Experimental Psychology: Applied, 14*(4), 392–400.

Fröhlich, P., Baldauf, M., Hagen, M., Suette, S., Schabus, D., & Kun, A. L. (2011). *Investigating Safety Services on the Motorway: The Role of Realistic Visualization*. Paper presented at the 3rd International Conference on Automotive User Interfaces and Interactive Vehicular Applications.

Gabbard, J. L., Fitch, G. M., & Kim, H. (2014). Behind the glass: Driver challenges and opportunities for AR automotive applications. *Proceedings of the IEEE, 102*(2), 124–136.

Gaspar, J. G., Street, W. N., Windsor, M. B., Carbonari, R., Kaczmarski, H., Kramer, A. F., & Mathewson, K. E. (2014). Providing views of the driving scene to drivers' conversation partners mitigates cell-phone-related distraction. *Psychological Science, 25*(12), 2136–2146.

Grandhi, S. A., & Jones, Q. (2015). Knock, knock! Who's there? Putting the user in control of managing interruptions. *International Journal of Human-Computer Studies, 79*, 35–50.

Gray, W. D., Sims, C. R., Fu, W.-T., & Schoelles, M. J. (2006). The soft constraints hypothesis: A rational analysis approach to resource allocation for interactive behavior. *Psychological Review, 113*(3), 461–482.

Heeman, P. A., Meshorer, T., Kun, A. L., Palinko, O., & Medenica, Z. (2013). *Estimating Cognitive Load Using Pupil Diameter during a Spoken Dialogue Task*. Paper presented at the 5th International Conference on Automotive User Interfaces and Interactive Vehicular Applications.

Iqbal, S. T., Horvitz, E., Ju, Y.-C., & Matthews, E. (2011). Hang on sec! Effects of proactive mediation of phone conversations while driving. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 463–472.

Iqbal, S. T., Ju, Y.-C., & Horvitz, E. (2010). *Cars, Calls, and Cognition: Investigating Driving and Divided Attention.*

Paper presented at the SIGCHI Conference on Human Factors in Computing Systems.

Janssen, C. P., Iqbal, S. T., & Ju, Y.-C. (2014). Sharing a driver's context with a caller via continuous audio cues to increase awareness about driver state. *Journal of Experimental Psychology: Applied, 20*(3), 270.

Klauer, S. G., Guo, F., Simons-Morton, B. G., Ouimet, M. C., Lee, S. E., & Dingus, T. A. (2014). Distracted driving and risk of road crashes among novice and experienced drivers. *New England Journal of Medicine, 370*(1), 54–59. Retrieved from http://doi.org/10.1056/NEJMsa1204142

Kun, A. L., Boll, S., & Schmidt, A. (2016). Shifting Gears: User interfaces in the age of autonomous driving. *IEEE Pervasive Computing, 15*(1), 32–38.

Kun, A. L., & Medenica, Z. (2012). *Video Call, or Not, That Is the Question*. Paper presented at the SIGCHI Conference on Human Factors in Computing Systems Extended Abstracts.

Kun, A. L., Palinko, O., Medenica, Z., & Heeman, P. A. (2013). *On the Feasibility of Using Pupil Diameter to Estimate Cognitive Load Changes for In-Vehicle Spoken Dialogues*. Paper presented at Interspeech.

Kun, A. L., Shyrokov, A., & Heeman, P. A. (2010). *Spoken Tasks for Human-Human Experiments: Towards In-Car Speech User Interfaces for Multi-Threaded Dialogue*. Paper presented at the 2nd International Conference on Automotive User Interfaces and Interactive Vehicular Applications.

Kun, A. L., Shyrokov, A., & Heeman, P. A. (2013). Interactions between human–human multi-threaded dialogues and driving. *Personal and Ubiquitous Computing, 17*(5), 825–834.

Kun, A. L., van der Meulen, H., & Janssen, C. P. (2017). *Calling while Driving: An Initial Experiment with Hololens*. Paper presented at the Driving Assessment 2017.

Kun, A. L., Wachtel, J., Miller, W. T., Son, P., & Lavallière, M. (2015). *User Interfaces for First Responder Vehicles: Views from Practitioners, Industry, and Academia*. Paper presented at the 7th International Conference on Automotive User Interfaces and Interactive Vehicular Applications.

Kunar, M. A., Carter, R., Cohen, M., & Horowitz, T. S. (2008). Telephone conversation impairs sustained visual attention via a central bottleneck. *Psychonomic Bulletin & Review, 15*(6), 1135–1140.

Maciej, J., Nitsch, M., & Vollrath, M. (2011). Conversing while driving: The importance of visual information for conversation modulation. *Transportation Research Part F: Traffic Psychology and Behaviour, 14*(6), 512–524.

McGill, M., Ng, A., & Brewster, S. (2017). *I Am the Passenger: How Visual Motion Cues Can Influence Sickness for In-Car VR*. Paper presented at the SIGCHI Conference on Human Factors in Computing Systems.

Medenica, Z., Kun, A. L., Paek, T., & Palinko, O. (2011). *Augmented Reality vs. Street Views: A Driving Simulator Study Comparing Two Emerging Navigation Aids*. Paper presented at the Proceedings of the 13th International Conference on Human Computer Interaction with Mobile Devices and Services.

Milgram, P., & Fumio K. (1994). A taxonomy of mixed reality visual displays. *IEICE Transactions on Information and Systems, 77*(12), 1321–1329.

Neale, V. L., Dingus, T. A., Klauer, S. G., Sudweeks, J., & Goodman, M. (2005). Overview of the 100-Car Naturalistic Study and Findings. National Highway Traffic Safety Administration, Paper (05-0400).

Pfleging, B., Schneegass, S., & Schmidt, A. (2013). *Exploring User Expectations for Context and Road Video Sharing while Calling and Driving*. Paper presented at the 5th International Conference on Automotive User Interfaces and Interactive Vehicular Applications.

SAE. (2015). Operational Definitions of Driving Performance Measures and Statistics (Recommended practice J2944). Warrendale, PA: Society of Automotive Engineers.

Salvucci, D. D. (2001). Predicting the effects of in-car interface use on driver performance: An integrated model approach. *International Journal of Human-Computer Studies, 55*(1), 85–107.

Schneider, M., & Kiesler, S. (2005). *Calling while Driving: Effects of Providing Remote Traffic Context*. Paper presented at the SIGCHI Conference on Human Factors in Computing Systems.

Senders, J. W., Kristofferson, A., Levison, W., Dietrich, C., & Ward, J. (1967). The attentional demand of automobile driving. *Highway Research Record* (195).

Smith, M., Gabbard, J. L., Burnett, G., & Doutcheva, N. (2017). The effects of augmented reality head-up displays on drivers' eye scan patterns, performance, and perceptions. *International Journal of Mobile Human Computer Interaction, 9*(2), 1–17.

Strayer, D. L., & Johnston, W. A. (2001). Driven to distraction: Dual-task studies of simulated driving and conversing on a cellular telephone. *Psychological Science, 12*(6), 462–466.

van der Heiden, R. M. A., Janssen, C. P., Donker, S. F., & Merkx, C. L. (2018, online first). Visual in-car warnings: How fast do drivers respond? *Transportation Research Part F: Traffic Psychology and Behaviour*.

Wintersberger, P., Sawitzky, T. V., Frison, A.-K., & Riener, A. (2017). *Traffic Augmentation as a Means to Increase Trust in Automated Driving Systems*. Paper presented at the 12th Biannual Conference on Italian SIGCHI Chapter, Cagliari, Italy.

Xiao, R., & Benko, H. (2016). *Augmenting the Field-of-View of Head-Mounted Displays with Sparse Peripheral Displays*. Paper presented at the SIGCHI Conference on Human Factors in Computing Systems.